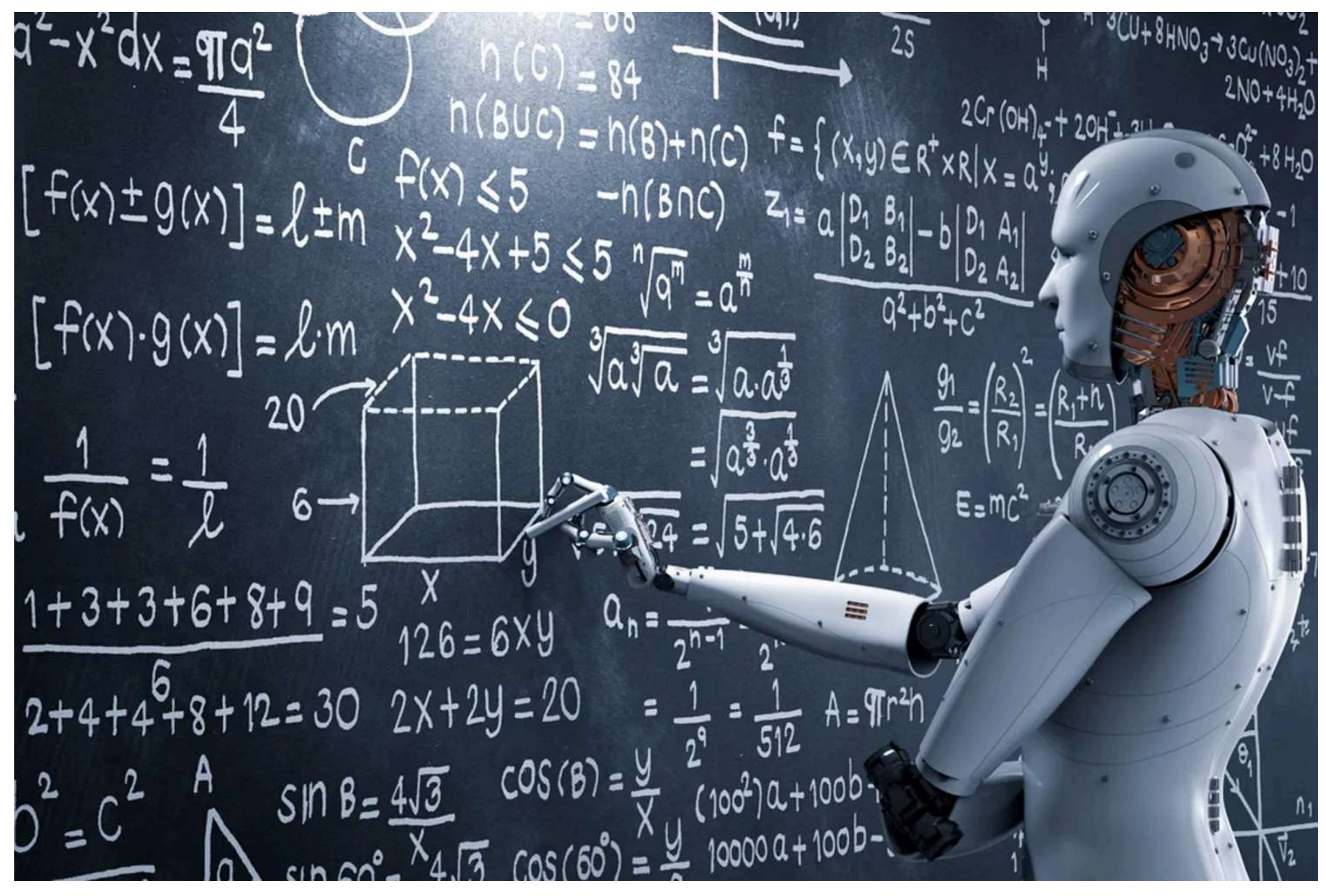
تكنولوجيا

جامعة «جونز هوبكنز»: نماذج الذكاء الاصطناعي تفشل في فهم التفاعلات البشرية



أظهرت الدراسة أن البشر يتفقون بدرجة عالية على تقييم مشاهد التفاعل الاجتماعي في حين فشل أكثر من 350 نموذجاً للذكاء الاصطناعي في محاكاتهم

لندن: نسيم رمضان

نُشر: 17:00-25 أبريل 2025 م . 27 شوّال 1446 هـ

في سباق تطوير أنظمة ذكاء اصطناعي أكثر ذكاءً وتفاعلاً مع البشر، تكشف دراسة جديدة من جامعة «جونز هوبكنز» حقيقة مقلقة؛ أنه حتى أكثر تقنيات الذكاء الاصطناعي تقدماً لا تزال تعاني من فهم أحد أبسط الأمور التي يجيدها البشر بالفطرة، وهو تفسير التفاعلات الاجتماعية.

من السيارات ذاتية القيادة إلى الروبوتات المساعدة والمساعدين الافتراضيين، يعتمد مستقبل الذكاء الاصطناعي على قدرته على قراءة السلوك البشري في السياقات الحقيقية. لكن وفقاً لهذه الدراسة الرائدة، لا يزال الذكاء الاصطناعي غير قادر على «قراءة المشهد».

تقول ليلى إيشيك، الأستاذة المساعدة في علم الإدراك بجامعة جونز هوبكنز، والمؤلفة الرئيسية للدراسة: «إن هذه الأنظمة تعاني من قيود أساسية عندما يتعلق الأمر بفهم كيفية تفاعل البشر بعضهم مع بعض». وتضيف: «إذا كنت تصمم ذكاءً اصطناعياً لسيارة ذاتية القيادة، فأنت بحاجة إلى أن يتنبأ بحركة المشاة. كأن يعرف ما إذا كان أحدهم على وشك العبور، أو ما إذا كان شخصان يتحدثان فقط. في الوقت الحالي، الذكاء الاصطناعي لا يستطيع القيام بذلك بدقة كافية».



دراسة «جونز هوبكنز»: لا تزال نماذج الذكاء الاصطناعي رغم تطورها غير قادرة على فهم التفاعلات الاجتماعية كما يفعل البشر

تطبيق عملي

تم عرض الدراسة في مؤتمر التعلم التمثيلي الدولي (ICLR) حيث حاول الباحثون اختبار قدرة الذكاء الاصطناعي على فهم الديناميكيات الاجتماعية، مقارنة بالبشر. صمّم الفريق تجربة بسيطة، لكنها فعّالة، فطُلب من المشاركين البشريين مشاهدة مقاطع فيديو قصيرة، مدتها 3 ثوانٍ لأشخاص، إما يتفاعل بعضهم مع بعض، أو يؤدون أنشطة متجاورة، أو يعملون بشكل مستقل. بعد ذلك، طُلب منهم تقييم مدى التفاعل الاجتماعي في كل مقطع على مقياس من 1 إلى 5.

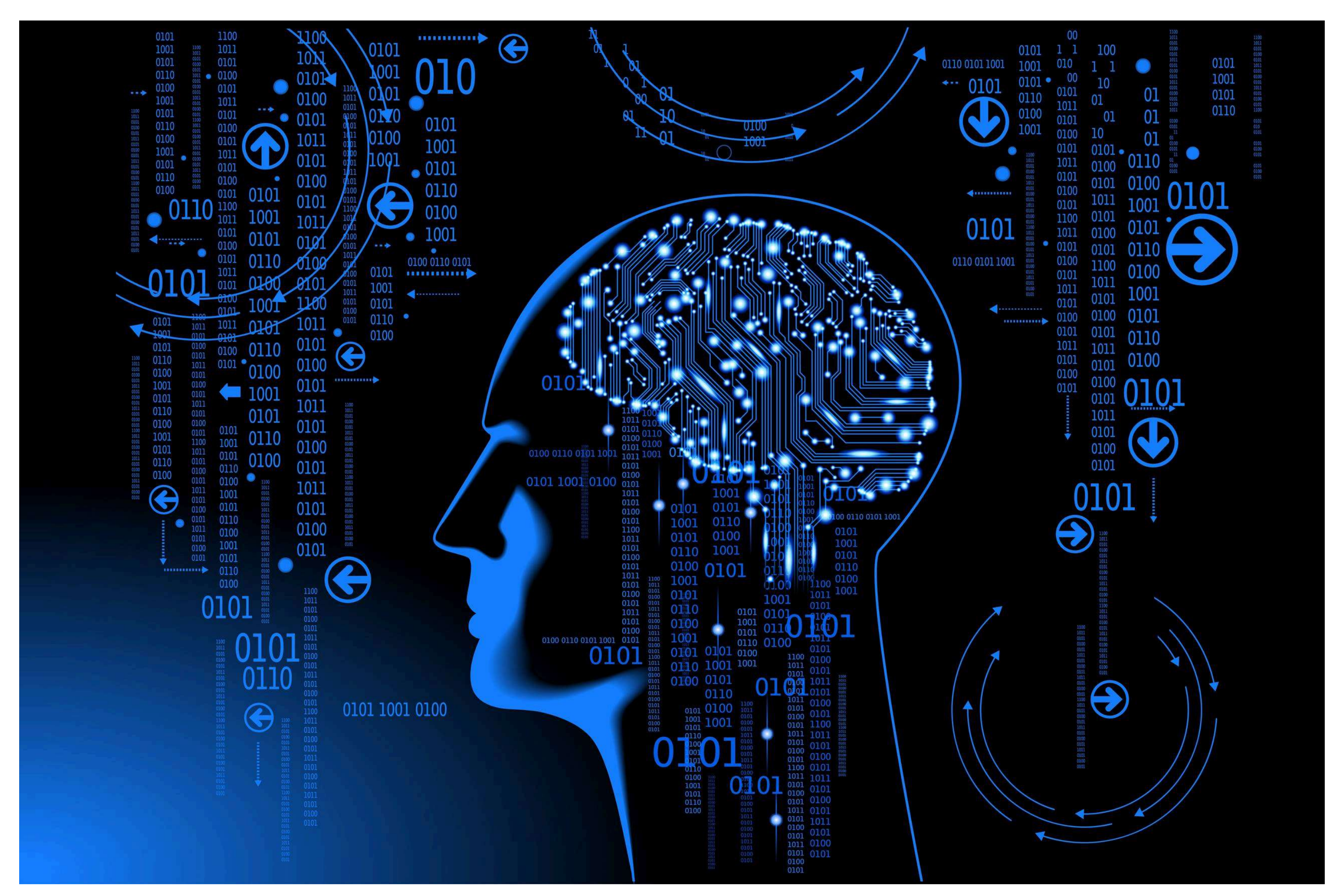
ثم طُبّق التقييم نفسه على أكثر من 350 نموذج ذكاء اصطناعي، بما في ذلك نماذج متقدمة لمعالجة الصور والفيديو واللغة. وكانت النتائج واضحة.

نتائج مقلقة

اتفق المشاركون من البشر إلى حدّ كبير في تقييماتهم، لكن نماذج الذكاء الاصطناعي بغضّ النظر عن مدى تطورها أو حجم بياناتها فشلت في الوصول إلى نفس الدقة. النماذج المعتمدة على الفيديو كانت الأسوأ أداءً، إذ لم تتمكن من التعرف على التفاعلات الاجتماعية في المشاهد الديناميكية. حتى النماذج المعتمدة على الصور الثابتة لم تستطع بدقة التنبؤ بما إذا كان الأشخاص يتواصلون أم لا.

أما النماذج اللغوية، التي أُعطيت أوصافاً مكتوبة للمشاهد، فقد أظهرت أداءً أفضل إلى حد ما، خصوصاً في محاولة التنبؤ بكيفية تفسير البشر للمشاهد. لكنها لم تتمكن من التقاط الصورة الكاملة لسياق المشهد المرئي المتغير. وترى كاتي غارسيا، الباحثة في مختبر إيشيك والمؤلفة المشاركة الأولى، التي عرضت النتائج في المؤتمر، أن هذا يعكس قصوراً بنيوياً في الطريقة التي تُبنى بها أنظمة الذكاء الاصطناعي.

تقول: «إن القدرة على التعرف على الوجوه أو الأجسام في الصور الثابتة كانت إنجازاً مهماً في مسار الذكاء الاصطناعي، لكنها ليست كافية. الحياة ليست صورة جامدة. الناس يتحركون، يتفاعلون، يتجاوب بعضهم مع بعض. وهذا المستوى من الفهم هو ما نحتاج من الذكاء الاصطناعي أن يصل إليه. ودراستنا تظهر أننا لم نصل بعد».



سبب القصور يعود إلى أن بنية الشبكات العصبية للذكاء الاصطناعي مستوحاة من مناطق الدماغ التي تعالج الصور الثابتة وليس التفاعلات الاجتماعية المعقدة (شاترستوك)

ما السبب؟

يعتقد الباحثون أن السبب يعود إلى الأساس الذي بُنيت عليه الشبكات العصبية للذكاء الاصطناعي، وهو جزء الدماغ المسؤول عن معالجة الصور الثابتة. أما التفاعلات الاجتماعية، فهي تتطلب فهماً أعمق للحركة والسياق والنية، وهي عمليات يعالجها جزء مختلف تماماً في دماغ الإنسان. توضح إيشيك أن هناك تفاصيل كثيرة، لكن الرسالة الكبرى هي أن أياً من نماذج الذكاء الاصطناعي التي اختبرناها لم تتمكن من مطابقة استجابات الدماغ البشري أو السلوك الإنساني لمشاهد متحركة، كما تفعل مع الصور الثابتة. هناك شيء أساسي في طريقة معالجة البشر للمشاهد الديناميكية، لا تزال هذه النماذج تفتقده.

تأثيرات بعيدة المدى

مع دخول الذكاء الاصطناعي إلى حياتنا اليومية، سواء في السيارات أو المستشفيات حتى الأماكن العامة، يصبح من الضروري أن يفهم سلوك البشر في السياق الاجتماعي بشكل دقيق. تردد بسيط من أحد المشاة، لغة جسد خفية، حتى تبادل نظرات، كلها إشارات يفسرها الإنسان دون تفكير، لكنها لا تزال لغزاً للذكاء الاصطناعي. هذه الدراسة لا تدعو إلى إيقاف استخدام الذكاء الاصطناعي في هذه السياقات، بل تؤكد على الحاجة إلى إعادة التفكير في كيفية تدريب وتصميم هذه الأنظمة، لتتشابه أكثر مع طريقة تفكير الدماغ البشري.

تختتم غارسيا قائلة إن هذه دعوة للاستيقاظ. وتضيف: «إذا أردنا أن نبني ذكاءً اصطناعياً يفهمنا فعلاً، فعلينا أن نعيد النظر في الطريقة التي نصمم بها هذه النماذج، من الأساس».

في نهاية المطاف، ربما يكون أحد أكبر تحديات الذكاء الاصطناعي ليس الفوز على البشر في الألعاب أو تأليف القصائد، بل أن يفهم نظرة أو إيماءة أو لحظة صامتة بين اثنين.

وعلى الأقل حتى الآن، يبقى البشر هم الأفضل في ذلك.

مواضيع تقنية | تكنولوجيا | الذكاء الاصطناعي | دراسة | أميركا | السعودية